IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application

Applicant(s): Drissi et al.
Docket No.: YOR920000401US1
Serial No.: 09/713,342
Filing Date: November 14, 2000
Group: 2121
Examiner: Wilbert L. Starks

I hereby certify that this paper is being deposited on this date with the U.S. Postal Service as first class mail addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450

Signature: _____ Date: November 17, 2004

Title: Method and Apparatus for Generating a Data Classification Model Using an Adaptive Learning Algorithm

---

APPEAL BRIEF

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Applicants hereby appeal the final rejection dated June 22, 2004, of claims 1 through 23 of the above-identified patent application.

REAL PARTY IN INTEREST

The present application is assigned to International Business Machines Corporation, as evidenced by an assignment recorded on November 14, 2000 in the United States Patent and Trademark Office at Reel 011307, Frame 0774. The assignee, International Business Machines Corporation, is the real party in interest.

RELATED APPEALS AND INTERFERENCES

There are no related appeals or interferences.

STATUS OF CLAIMS

Claims 1 through 23 are pending in the above-identified patent application. Claims 1-23 remain rejected as being directed to non-statutory subject

matter. Claims 8, 9, 21, and 23 remain rejected under 35 U.S.C. § 102(b) as being anticipated by McAulay, A.D. and Oh, J.C., Improved Learning in Genetic Rule-Based Classifier Systems, Systems, Man and Cybernetics, 1991; Decision Aiding for Complex Systems, Conference Proceedings, 1991 IEEE International Conference, October 13-16, 1991, Pages 1393-1398, Vol. 2 (hereinafter McAulay), and claims 1-23 remain rejected under 35 U.S.C. §103(a) as being unpatentable over McAulay et al. in view of Lewis, David D., An Evaluation of Phrasal and Clustered Representations on a Text Categorization Task, Proceedings of the Fifteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, June 1992, pages 37-50 (hereinafter Lewis).

STATUS OF AMENDMENTS

There have been no amendments filed subsequent to the final rejection.

SUMMARY OF INVENTION

The present invention is directed to a data classification method and apparatus for labeling unknown objects. The disclosed data classification system employs a learning algorithm that adapts through experience. The present invention classifies objects in domain datasets using data classification models having a corresponding bias and evaluates the performance of the data classification. The performance values for each domain dataset and corresponding model bias are processed to identify or modify one or more rules of experience. (Page 9, line 4, to page 10, line 3.) The rules of experience are subsequently used to generate a model for data classification. Each rule of experience specifies one or more characteristics for a domain dataset and a corresponding bias that should be utilized for a data classification model if the rule is satisfied. (Page 10, lines 4-24.) The present invention dynamically modifies the assumptions (bias) of the learning algorithm to improve the assumptions embodied in the generated models and thereby improve the quality of the data classification and regression systems that employ such models. The disclosed self-adaptive learning process will become increasingly more accurate as the rules of experience are accumulated over time. (Page 10, line 25, to page 11, line 18.)

ISSUES PRESENTED FOR REVIEW

i. Whether claims 1-23 are properly rejected as being directed to non-statutory subject matter;

ii. whether claims 8, 9, 21, and 23 are properly rejected under 35 U.S.C. § 102(b) as being anticipated by McAulay; and

iii. whether claims 1-23 are properly rejected under 35 U.S.C. §103(a) as being unpatentable over McAulay et al. in view of Lewis.

GROUPING OF CLAIMS

The rejected claims do not stand and fall together. More particularly, for the reasons given below, Applicants believe that each of the dependent claims 3/18 and 4/19 provide independent bases for patentability apart from the rejected independent claims.

ARGUMENT

Section 101 Rejections

Claims 1-23 were rejected as being directed to non-statutory subject matter. In particular, the Examiner asserts that claims 1, 8, 13, 16, and 21-23 are not claimed to be practiced on a computer and that it is clear that these claims are not limited to practice in the technological arts. The Examiner further asserts that none of the claims are limited to practical applications in the technological arts, that Applicants fail to define a useful, concrete and tangible result, and do not specify the associated practical application with the appropriate level of specificity. The Examiner also finds that the Applicants manipulated a set of abstract "input data" to solve mathematical problems in the abstract and that the result of such manipulations is not statutory. Regarding the "system" and "computer readable medium" recitals in claims 16-23, the Examiner asserts that the invention is still found to be non-statutory.

Under Section 101, "any new and useful process, machine, manufacture, or composition of matter" is patentable. 35 U.S.C. §101. It is recognized, however, that despite the broad scope of section 101, "laws of nature, physical phenomena and abstract

ideas" cannot be patented. *Diamond v. Chakrabarty*, 447 U.S. 303, 309, 206 U.S.P.Q. (BNA) 193, 197 (1980).

The Examiner asserts that Claims 1-23 are not claimed to be practiced on a computer and that it is clear that these claims are not limited to practice in the technological arts. To the contrary, however, each of the independent claims are expressly directed to a practical method of (or system for) "classifying data." For example, the method can be used to classify real numerical vectors. Thus, each of these claims are clearly tied to a practical application. A process that is limited to a practical application of an abstract idea or mathematical algorithm in the technological arts is patentable. *See* Examination Guidelines for Computer-Related Inventions, Section IV. B. 2. b. (ii).

In any event, the analysis does not stop there. The Supreme Court has stated that the "***[t]ransformation*** and reduction of an article 'to a different state or thing' is the clue to patentability of a process claim." *Gottshalk v. Benson*, 409 U.S. 63, 70, 175 U.S.P.Q. (BNA) 676 (1972). In other words, claims that require some kind of transformation of subject matter, which has been held to include intangible subject matter, such as data or signals that are representative of or constitute physical activity or objects, have been held to comply with Section 101. *See, for example, In re Warmerdam,* 31 U.S.P.Q.2d (BNA) 1754, 1759 n.5 (Fed. Cir. 1994) or *In re Schrader*, 22 F.3d 290, 295, 30 U.S.P.Q.2d (BNA) 1455, 1459 n.12 (Fed. Cir. 1994).

Each independent claim includes at least one transformation. For example, independent claims 1, 16 and 22 **modify** the bias of one or more data classification models, based on a performance evaluation. Thus, a modified data classification model is provided. Claims 8, 21 and 23 **classify** objects and **select** a data classification model for classifying a domain dataset by comparing characteristics of the domain dataset to rules. Thus, an object classification is provided. Finally, claim 13 processes performance values for each combination of domain dataset and said bias to **adjust** one or more rules for subsequent data classification. Thus, adjusted rules are provided.

Applicants submit that each of the claims 1-23 are in full compliance with 35 U.S.C. §101, and accordingly, respectfully request that the rejection under 35 U.S.C. §101 be withdrawn.

Independent Claims 1, 8, 13, 16 and 21-23

Independent claims 8, 9, 21, and 23 are rejected under 35 U.S.C. § 102(b) as being anticipated by McAulay and independent claims 1, 8, 13, 16, and 21-23 are rejected under 35 U.S.C. §103(a) as being unpatentable over McAulay et al. in view of Lewis.

Regarding claim 1, the Examiner acknowledges that McAulay does not disclose selecting at least one of said one or more data classification models based on a meta-feature that characterizes said domain data set, but asserts that Lewis does show a classifier using meta-features. Regarding claims 8, 21, and 23, the Examiner asserts that McAulay teaches selecting a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules (FIG. 1: lines 4-5).

Regarding claim 1, Applicants note that Lewis teaches that "most current indexing languages represent documents as tuples or vectors of numeric or binary values, with *each value corresponding to an indexing term*." (Page 38, Section 2.) Lewis then teaches that, "for clarity, we therefore call the features of indexing terms metafeatures." (Page 38, Section 2.2). *Metafeatures in Lewis are therefore features of indexing terms* (the individual values representing a document) and not domain datasets. More importantly, Lewis does not disclose selecting data classification models based on a *meta-feature that characterizes a domain data set*. In addition, since Lewis only discloses the use of one algorithm (the genetic algorithm), there is no *selection of classification models*. Independent claims 1, 16, and 22 require classifying objects in a domain dataset using one or more data classification models, each of said one or more data classification models having a bias; selecting at least one of said one or more data classification models based on a meta-feature that characterizes said domain data set; evaluating the performance of said classifying step; and modifying said bias based on said performance evaluation. Independent claim 13 requires applying an adaptive learning algorithm to said domain dataset to select a data classification model based on a meta-feature that characterizes said domain data set, said data classification model having

a bias; classifying objects in said domain dataset using said selected data classification model; evaluating the performance of said classifying step; maintaining an indication of said performance of said model for said domain dataset; repeating said applying, classifying and evaluating steps for a plurality of said domain datasets; and processing said performance values for each combination of said domain datasets and said bias to adjust one or more rules for subsequent data classification, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models. Independent claim 8, 21, and 23 require classifying objects in a plurality of domain datasets using one of a number of data classification models, each of said data classification models having a corresponding bias; evaluating the performance of each of said domain dataset classifications; maintaining a performance value for each combination of said domain datasets and said bias; processing said performance values for each combination of said domain datasets and said bias to generate one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models; and selecting a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules.

Thus, McAulay et al. or Lewis, alone or in combination, do not disclose or suggest classifying objects in a domain dataset using one or more data classification models, each of said one or more data classification models having a bias; selecting at least one of said one or more data classification models based on a meta-feature that characterizes said domain data set; evaluating the performance of said classifying step; and modifying said bias based on said performance evaluation, as required by independent claims 1, 16, and 22, do not disclose or suggest applying an adaptive learning algorithm to said domain dataset to select a data classification model based on a meta-feature that characterizes said domain data set, said data classification model having a bias; classifying objects in said domain dataset using said selected data classification model; evaluating the performance of said classifying step; maintaining an indication of said performance of said model for said domain dataset; repeating said applying, classifying and evaluating steps for a plurality of said domain datasets; and processing

said performance values for each combination of said domain datasets and said bias to adjust one or more rules for subsequent data classification, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models, as required by independent claim 13, and do not disclose or suggest classifying objects in a plurality of domain datasets using one of a number of data classification models, each of said data classification models having a corresponding bias; evaluating the performance of each of said domain dataset classifications; maintaining a performance value for each combination of said domain datasets and said bias; processing said performance values for each combination of said domain datasets and said bias to generate one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models; and selecting a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules, as required by independent claims 8, 21, and 23.

Conclusion

The rejections of the independent claims under §102 and §103 in view of McAulay et al. or Lewis, alone or in any combination, are therefore believed to be improper and should be withdrawn.

Dependent Claims

Claims 3/18 and 4/19 specify a number of limitations providing additional bases for patentability. Specifically, the Examiner rejected claims 3, 4, 18, and 19 under 35 U.S.C. §103(a) as being unpatentable over McAulay et al. in view of Lewis. Claims 3 and 18 require the step of processing said recorded performance values for each combination of said domain datasets and said bias to generate one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models. Claims 4 and 19 require the step of selecting a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules.

The Examiner asserts that the limitation of claim 3 is taught by McAulay (FIG. 1: lines 4-5). Applicants note, however, that McAulay does not disclose or suggest generating one or more rules, *each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models*, as required by dependent claims 3 and 18.

The Examiner asserts that the limitation of claim 4 is taught by McAulay (Page 1393, third paragraph, first three lines of the paragraph). Applicants note, however, that McAulay does not disclose or suggest the step of selecting a data classification model for *classifying a domain dataset by comparing characteristics of said domain dataset to said rules*, as required by dependent claims 4 and 19.

Thus, McAulay et al. or Lewis, alone or in combination, do not disclose or suggest generating one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models, as required by dependent claims 3 and 18, and do not disclose or suggest the step of selecting a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules, as required by dependent claims 4 and 19.

The remaining rejected dependent claims are believed allowable for at least the reasons identified above with respect to the independent claims.

The attention of the Examiner and the Appeal Board to this matter is appreciated.

Respectfully,

Kevin M. Mason

Date: November 17, 2004

Kevin M. Mason
Attorney for Applicant(s)
Reg. No. 36,597
Ryan, Mason & Lewis, LLP
1300 Post Road, Suite 205
Fairfield, CT 06824
(203) 255-6560

APPENDIX

1. A method for classifying data, comprising the steps of:

classifying objects in a domain dataset using one or more data classification models, each of said one or more data classification models having a bias;

selecting at least one of said one or more data classification models based on a meta-feature that characterizes said domain data set;

evaluating the performance of said classifying step; and

modifying said bias based on said performance evaluation.

2. The method of claim 1, wherein said steps of classifying and evaluating are performed for a plurality of said domain datasets and wherein said method further comprising the steps of recording a performance value for each combination of said domain datasets and said bias.

3. The method of claim 2, further comprising the step of processing said recorded performance values for each combination of said domain datasets and said bias to generate one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models.

4. The method of claim 3, further comprising the step of selecting a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules.

5. The method of claim 1, wherein said domain dataset is represented using a set of meta-features.

6. The method of claim 5, wherein said meta-features includes a concept variation meta-feature.

7. The method of claim 5, wherein said meta-features includes an average weighted distance meta-feature that measures the density of the distribution of said at least one domain dataset.

8. A method for classifying data, comprising the steps of:

classifying objects in a plurality of domain datasets using one of a number of data classification models, each of said data classification models having a corresponding bias;

evaluating the performance of each of said domain dataset classifications;

maintaining a performance value for each combination of said domain datasets and said bias;

processing said performance values for each combination of said domain datasets and said bias to generate one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models; and

selecting a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules.

9. The method of claim 8, further comprising the step of modifying at least one of said biases based on said performance evaluation.

10. The method of claim 8, wherein said domain dataset is represented using a set of meta-features.

11. The method of claim 10, wherein said meta-features includes a concept variation meta-feature.

12. The method of claim 10, wherein said meta-features includes an average weighted distance meta-feature that measures the density of the distribution of said at least one domain dataset.

13. A method for classifying data in a domain dataset, comprising:

applying an adaptive learning algorithm to said domain dataset to select a data classification model based on a meta-feature that characterizes said domain data set, said data classification model having a bias;

classifying objects in said domain dataset using said selected data classification model;

evaluating the performance of said classifying step;

maintaining an indication of said performance of said model for said domain dataset;

repeating said applying, classifying and evaluating steps for a plurality of said domain datasets; and

processing said performance values for each combination of said domain datasets and said bias to adjust one or more rules for subsequent data classification, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models.

14. The method of claim 13, further comprising the step of selecting a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules.

15. The method of claim 13, further comprising the step of modifying at least one of said biases based on said performance evaluation.

16. A system for classifying data, comprising:

a memory that stores computer-readable code; and

a processor operatively coupled to said memory, said processor configured to implement said computer-readable code, said computer-readable code configured to:

classify objects in a domain dataset using ~~a~~ one or more data classification models, each of said one or more data classification models having a bias;

selecting at least one of said one or more data classification models based on a meta-feature that characterizes said domain data set;

evaluate the performance of said classifying step; and

modify said bias based on said performance evaluation.

17. The system of claim 16, wherein said processor is further configured to classify said objects and evaluate said performance for a plurality of said domain datasets and wherein said processor records a performance value for each combination of said domain datasets and said bias.

18. The system of claim 17, wherein said processor is further configured to process said recorded performance values for each combination of said domain datasets and said bias to generate one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models.

19. The system of claim 18, wherein said processor is further configured to select a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules.

20. The system of claim 16, wherein said domain dataset is represented using a set of meta-features.

21. A system for classifying data, comprising:

a memory that stores computer-readable code; and

a processor operatively coupled to said memory, said processor configured to implement said computer-readable code, said computer-readable code configured to:

classify objects in a plurality of domain datasets using one of a number of data classification models, each of said data classification models having a corresponding bias;

evaluate the performance of each of said domain dataset classifications;

maintaining a performance value for each combination of said domain datasets and said bias;

process said performance values for each combination of said domain datasets and said bias to generate one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models; and

select a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules.

22. An article of manufacture for classifying data, comprising:

a computer readable medium having computer readable code means embodied thereon, said computer readable program code means comprising:

a step to classify objects in a domain dataset using ~~a~~ one or more data classification models, each of said one or more data classification models having a bias;

selecting at least one of said one or more data classification models based on a meta-feature that characterizes said domain data set;

a step to evaluate the performance of said classifying step; and

a step to modify said bias based on said performance evaluation.

23. An article of manufacture for classifying data, comprising:

a computer readable medium having computer readable code means embodied thereon, said computer readable program code means comprising:

a step to classify objects in a plurality of domain datasets using one of a number of data classification models, each of said data classification models having a corresponding bias;

a step to evaluate the performance of each of said domain dataset classifications;

a step to maintaining a performance value for each combination of said domain datasets and said bias;

a step to process said performance values for each combination of said domain datasets and said bias to generate one or more rules, each of said rules specifying one or more characteristics of said domain datasets and a corresponding bias that should be utilized in one of said data classification models; and

a step to select a data classification model for classifying a domain dataset by comparing characteristics of said domain dataset to said rules.